Volume Title: The Smoothing of Time Series

Volume Author/Editor: Frederick R. Macaulay

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/maca31-1

Publication Date: 1931

Chapter Title: Dr. E. C. Rhodes' Graduation Proposal

Chapter Author: Frederick R. Macaulay

Chapter URL: http://www.nber.org/chapters/c9365

Chapter pages in book: (p. 81 - 88)

CHAPTER V

Dr. E. C. Rhodes' Graduation Proposal.

In 1921, Dr. E. C. Rhodes proposed a method of smoothing based on the fitting of fourth-degree parabolas.[1] The method is somewhat different from any other discussed in this book.

The graduation is accomplished as follows: First, from an examination of the data, decide upon a certain number of terms—for example, 15. Then, to the first 15 observations fit a fourth-degree parabola by the method of least squares. Take the first 8 points on this parabola as the first 8 points of the desired smooth curve. To the second 15 observations (points 2 to 16, inclusive) fit a fourth-degree parabola by the method of least squares in such a manner that this second parabola passes through the 8th point of the first parabola and at that 8th point has the same slope as the first parabola. Obtain the 9th point on the smooth curve from this second parabola. Repeat the operation for each successive point on the smooth curve until such a "conditioned" parabola has been fitted to the last 15 points of the data. Use the

[1] E. C. Rhodes, *Smoothing*—Tracts for Computers, No. 6. Edited by Karl Pearson, Cambridge University Press, 1921.

last 8 points on this parabola as the last 8 points of the smooth curve.

This method of fitting can evidently give some strange results, as may be seen from an examination of the tables and charts in Dr. Rhodes' monograph. The closeness of fit of the graduation to the test data [1] is not connected in any simple manner with the number of terms to which the fourth-degree parabola is fitted. Decreasing the number of terms sometimes increases and sometimes decreases the goodness of fit. For example, Dr. Rhodes presents four graduations which he designates a 9-point curve, an 11-point curve, a 13-point curve, and a 15-point curve. A priori, one might think that, though the 9-point curve might not be as smooth as the 13 or the 15, it would lie closer to the data. Now, the sum of the squares of the deviations of the data from the fitted curve is much greater in the case of the 9-point curve than in the case of either the 13 or the 15. For the 29 observations which are covered by all four curves and which he uses as a comparison range, the sum of the squares of the deviations of the data from Dr. Rhodes' smooth curves are: for the

[1] Dr. Rhodes' test data are taken from an article by W. F. Sheppard entitled *Graduation by Reduction of Mean Square of Error*, Journal of the Institute of Actuaries, Vol. 48, p. 178. The data are rates of infantile mortality from causes other than diarrheal diseases, for the 42 years from 1870 to 1911, inclusive.

15-point curve, 266.0: for the 13-point curve, 262.2; for the 11-point curve, 243.6; and for the 9-point curve, 348.7.[1] In Dr. Rhodes' illustration, the 9-point curve is weirdly erratic, as may be seen from a glance at Diagram II at the end of his monograph.

Dr. Rhodes' fitted curves are smooth when the correct number of data items are taken in fitting his parabolas; otherwise strange, sweeping sinuosities are introduced which have little apparent relation to the data. For the 42 observations, the sum of the squares of the third differences of Rhodes' 9-point curve is 98.52. This compares with 11.75 for the 11-point curve, 5.58 for the 13-point curve and 1.61 for the 15-point curve. Dr. Rhodes was not pleased with the results he obtained by fitting the 9-point and 11-point curves. An examination of the table in Appendix V of this book, or Diagram II in the Appendix to Dr. Rhodes' book, will suggest the reason why. The sinuosities introduced have little apparent relation to the data. Dr. Rhodes was pleased with the results of his 13- and 15-point curves.

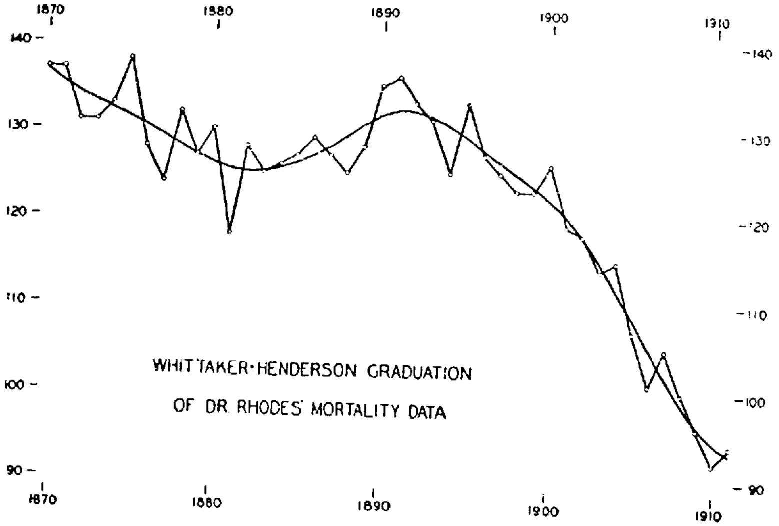
In Appendix V are given: I, Rhodes' data; II, Rhodes' 15-point curve; III, Rhodes' 13-point curve; IV, Rhodes' 11-point curve; V, Rhodes' 9-point curve; VI, Sheppard's graduation; VII,

[1] Rhodes, *Smoothing*, p. 25.

Spencer's 15-term graduation; VIII, a Whittaker-Henderson graduation with n = 2.[1]

Though the figures for the Spencer 15-term curve and the Whittaker-Henderson curve are given to two decimals in Appendix V, all the comparisons which follow are from calculations based on nearest first decimal places, as Dr. Rhodes' figures con-

CHART II



tain only one decimal. The first six columns of the table in Appendix V are from Dr. Rhodes' monograph. Columns VII and VIII were calculated in the National Bureau of Economic Research. Sheppard's graduation is not particularly interesting, though it seems to have been the point of departure for Dr. Rhodes' monograph. Spencer's

[1] For the Whittaker-Henderson method of graduation see Appendix VI.

15-term formula is attractive. The amount of labor involved is, even if the ends be extrapolated, only a small percentage of that necessary in the case of the Rhodes' method. The Whittaker-Henderson graduation gives the best results.

Some comparisons of the Whittaker-Henderson graduation and the Spencer 15-term graduation with the various results obtained by Dr. Rhodes' method may be interesting. The table below gives comparisons based upon the middle twenty-eight observations and the total forty-two observations. This classification is introduced in order to give a comparison of the different curves in the middle region where no extrapolation occurs, as well as for the whole range. The table consists of three parts. Part I compares the lack of smoothness of the various curves; Parts II and III give measures of their badness of fit.

MEASUREMENTS ON THE VARIOUS GRADUATIONS OF DR. RHODES' TEST DATA.

PART I.

Sums of the Squares of the Third Differences of the Various Curves.

	28 Observations	42 Observations
Rhodes' 15-point curve	1.10	1.61
" 13 " "	3.55	5.58
" 11 " "	7.95	11.75
" 9 " "	55.67	98.52
Dr. Sheppard's Graduation	39.45	45.12
Spencer's 15-term formula	1.46	2.42
Whittaker-Henderson ($n = 2$)	.77	1.15

PART II.

Sums of Deviations of Data from Curves.

(without regard to sign)

	28 Observations	42 Observations
Rhodes' 15-point curve	65.1	96.2
" 13 " "	65.9	97.3
" 11 " "	61.4	91.8
" 9 " "	69.9	109.7
Dr. Sheppard's Graduation............	68.1	99.7
Spencer's 15-term formula	64.0	93.9
Whittaker-Henderson (n = 2)	66.2	96.3

PART III.

Sums of the Squares of the Deviations of the Data from the Curves.

	28 Observations	42 Observations
Rhodes' 15-point curve	261.59	368.42
" 13 " "	257.77	361.67
" 11 " "	239.24	337.74
" 9 " "	345.07	488.85
Dr. Sheppard's Graduation	259.95	370.55
Spencer's 15-term formula	250.56	353.51
Whittaker-Henderson (n = 2)	256.70	359.17

There seems no good reason for substituting such a method as that of Dr. Rhodes for the much simpler summation method illustrated by Spencer's 15-term formula. While the results obtainable by the Whittaker-Henderson method are distinctly better than those obtainable by any one of the other methods, Dr. Rhodes cannot be criticized for neglecting this method, as it had not been published at the time he wrote his monograph.

Dr. Rhodes stresses the fact that his curve is a continuous curve,[1] each part of which may be represented by a mathematical equation, and that therefore interpolation or summation may be exactly accomplished. This is not highly important. Not only the Whittaker-Henderson curve, but also most good summation formulas give curves so smooth that parabolic interpolation between the points is simple and legitimate. Dr. Rhodes' method of graduating is not adapted to the elimination of seasonal fluctuations. It is also decidedly erratic in its results. Even with the short cuts Dr. Rhodes introduces, the calculation is laborious. We did not seriously consider using it in the interest rate and security price study.[2]

For a short series of data (such as the mortality figures used by Dr. Rhodes), which are not affected by seasonal fluctuations, a Whittaker-Henderson graduation is ideal.[3] From its very nature that method of smoothing gives perfect results if tested by the sum of the squares of the deviations

[1] Compare Rhodes, p. 8.

[2] The results obtained by Dr. Rhodes' method are slightly affected by the decision as to which end of the data shall be used as a beginning for operations. The summation formulas, and indeed all other formulas which are discussed in this book, give identical results whether they are worked forward or backward.

[3] Column VIII of Appendix V and Chart II give the results of applying such a graduation to Dr. Rhodes' test data. In this particular case n = 2. The significance of this statement will be understood when the immediately following text discussing the Whittaker-Henderson method of graduation has been read.

of the data from the graduation and the sum of the squares of the third differences of the graduation itself. When Henderson's method of computation is used, the labor involved in fitting to such a short series is not excessive. The sum of the squares of the 42 deviations of the data from the Whittaker-Henderson curve (with $n = 2$) is 359.17, which compares with 361.67 for the Rhodes 13-point curve, and 368.42 for the Rhodes 15-point curve. The sum of the squares of the third differences of the Whittaker-Henderson curve itself is 1.15, which compares with 1.61 for the Rhodes 15-point curve, and 5.58 for the Rhodes 13-point curve. The Whittaker-Henderson method of graduation is briefly explained in the immediately following chapter. The actual computation is outlined in Appendix VI.